Journal of Theoretical Biology 353 (2014) 55–60

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

A new index for the quantification of chromosome number variation: An application to selected animal and plant groups

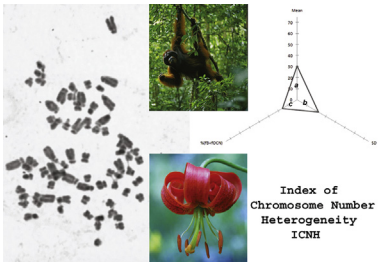
Lorenzo Peruzzi *, Katia Francesca Caparelli, Gianni Bedini

Dipartimento di Biologia, Unità di Botanica, Università di Pisa, Via Luca Ghini 13, 56126 Italy

HIGHLIGHTS

- The new Index of Chromosome Number Heterogeneity (ICNH) is proposed.
- ICNH is a useful parameter to compare related taxonomical/geographical groups of organisms.
- Significant differences among selected animal and plant groups are detected.
- Higher chromosome numbers are associated with a larger variation degree in plants, but not in animals.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 24 January 2014
Received in revised form
7 March 2014
Accepted 9 March 2014
Available online 16 March 2014

Keywords:
B chromosomes
Evolution
Index of Chromosome Number Heterogeneity
Odd chromosome numbers
Polyploidy

ABSTRACT

Quantitative parameters have been used to characterize chromosome number (CN) variation. This gave us the idea to collect available data in various organisms and compare them, in order to verify if variation patterns differ between animal and plant groups and to quantify these patterns with an Index of CN Heterogeneity (ICNH), useful as a parameter to compare related taxonomical/geographical groups of organisms. To the best of our knowledge, this is the first attempt to compare CN variation in animal and plant groups with large datasets. The quantitative analysis allowed detecting significant differences among most groups of animals and plants. The most striking difference, however, is the close relationship between mean CN and SD restricted to plants, in which higher CN are also associated with a larger variation degree, possibly due to the well known genomic plasticity in this group and a propensity for polyploidization higher than in animals. The ICNH defined here can be easily calculated for both animal and plant groups based on commonly available data. It summarizes data accumulated in over a century of research and includes so-called anomalies like *f*B and *f*OCN, sometimes overlooked by researchers due to lack of a proper way of comparison.

1. Introduction

The first chromosome counts were published in 1882 by Eduard Strasburger (1844–1912) and included several data concerning mostly some genera of vascular plants, but also salamanders (Garbari et al., 2012). This happened before the coining of the term chromosome itself ("*chromosom*") by Wilhelm Waldeyer in 1888 ("*If the term I propose is practically applicable it will become familiar, otherwise it will be soon sink into oblivion*" (Battaglia, 2009)) and well before the discovery of DNA as "giant hereditary molecule" in 1927 by Nikolai Koltsov (Soyfer, 2001).

Today, besides morphological data, chromosome number (CN hereafter) is the most widespread and cheap systematic character, while karyotype structure details, karyotype asymmetry, chromosome painting, genome size and molecular systematics have

List of abbreviations: CN, Chromosome number; *f*B, Frequency of B chromosomes; *f*OCN, Frequency of odd chromosome numbers; ICNH, Index of Chromosome Number Heterogeneity

* Corresponding author.
E-mail addresses: lorenzo.peruzzi@unipi.it (L. Peruzzi), kfcaparelli@gmail.com (K.F. Caparelli), gianni.bedini@unipi.it (G. Bedini).

http://dx.doi.org/10.1016/j.jtbi.2014.03.012
0022-5193/

a much narrower coverage in scientific literature (Bennett and Leitch, 2010; Garcia et al., 2012; Guerra, 2012; Siljak-Yakovlev and Peruzzi, 2012; Peruzzi and Eroğlu, 2013).

CN can vary among and within taxonomic groups of living organisms, and in more than 130 years a lot of data accumulated in literature, stimulating the onset of CN databases, either hard-printed (especially in the last century) or online.

Currently, the known CN range in organisms goes from $2n=2$ in the ant *Myrmecia pilosa* to $2n=1440$ in the psilophyte fern *Ophioglossum reticulatum* (Castiglione and Cremonini, 2012).

In recent years, researchers started to use existing CN databases (especially those online) not only as indexing tools for an easier retrieval of data (Bareka et al., 2008), but also as analytical tools, enabling to answer intriguing questions regarding the variation of CN in plants from different areas (Peruzzi et al., 2011, 2012), of different taxonomic groups (Graphodatsky et al., 2011; Bedini et al., 2011b) and between allochtonous and autochtonous taxa (Góralski et al., 2013).

Quantitative parameters have been used to characterize CN variation in a group: mean, median, standard deviation of CN, frequency of B chromosomes (Jones, 2012) and odd CNs (Peruzzi et al., 2011; Bedini et al., 2011b).

This gave us the idea to collect available data concerning CN in various organisms and compare them, in order to (a) verify if CN variation patterns differ between selected animal and plant groups; (b) quantify these patterns with an Index of CN Heterogeneity (ICNH), useful as a parameter to compare related taxonomical/geographical groups of organisms.

2. Materials and methods

A survey of recently published and/or online easily available web karyological resources allowed us to obtain data concerning 10 animal and 17 plant taxonomical/geographical groups, for a total of 5557 and 15,592 cytotypes, respectively (Table 1). The total number of cytotypes retained for each group was obtained by excluding counts in multiple copy (i.e. the same chromosome number for the same species). Any *n* count (a minority in the two datasets) was transformed to 2*n*.

For each group, the following parameters were calculated: (a) mean CN, (b) standard deviation of CN, (c) frequency of B chromosomes (*f*B hereafter); and (d) frequency of odd CN (*f*OCN). It must be noticed that B chromosomes were never counted as part of the A set of chromosomes of an organism.

To quantify the variation of chromosome number, we chose to calculate the square root of the area of the ideal triangle built in a three-variables radar plot, where the vertices of the triangle are defined by mean CN, SD of CN, and %(*f*B+*f*OCN). This triangle gives a graphical representation of CN variation in a group, and its area can be easily seen as the sum of the three areas subtended by the smaller triangles set along the plot axes (Fig. 1). We define here this value as the Index of Chromosome Number Heterogeneity (ICNH hereafter), which is easily calculated according to the following formula:

$$\mathrm{ICNH}=\sqrt{\sin\left(\frac{2\pi}{3}\right)\frac{ab+ac+bc}{2}}$$

where *a* is CN, *b* is SD of CN, *c* is %(*f*B+*f*OCN). Its value can vary from 0 (if no CN variation occurs in a group, i.e. that group would be represented by a vertical line only varying in length) to $+\infty$, albeit very high values can be reached only theoretically.

For multiple pairwise comparisons among characters, ANOVA (with Tukey HSD test) was used within R (http://www.r-project.org) to test the occurrence of statistically significant differences. If Levene statistics revealed a non-homoscedasticity of variance, then the non-parametric Kruskal–Wallis test for independent samples was used, with Bonferroni correction for multiple comparisons. The same software was used to calculate correlation coefficients among the considered parameters; correlations were considered as weak (up to 0.3), moderate (up to 0.7), strong (> 0.7) (Peruzzi et al., 2009). Only correlations significant at the 99% level or higher are presented and discussed.

Table 1
Animal and plant chromosome number resources used in this study.

	Reference
Animals	
Argentinian freshwater fishes	Fenocchio et al. (2003)
Cyclopoida	Yang et al. (2008)
Georgian earthworms	Bakhtadze et al. (2008)
Indian ants	Imai et al. (1984)
Mammalia	Graphodatsky et al. (2000)
Primates	Romagno (2001)
Pseudopimelodidae +Pimelodidae+Heptaptedidae	Swarça et al. (2007)
Psyllidae	Maryańska-Nadachowska (2002)
Reptilia	Olmo and Signorino (2012)
Salamanders	Sessions (2007)
plants	
Chilean vascular flora	Jara-Seguel and Urrutia (2011,2012)
Cyperaceae	Roalson (2008)
Gagea	Peruzzi (2003, 2008) and Peruzzi and Aquaro (2005)
Hieracium	Schuhwerk (2012)
Iridaceae	Goldblatt and Takei (1997)
Italian vascular flora	Bedini et al. (2010, 2012a)
Liliaceae+Smilacaceae	Peruzzi et al. (2009)
Malvaceae	Hinsley (2009)
Myrtaceae	Rye (1979)
New Zealand mosses	Ramsay (2009)
New Zealand vascular flora	Peruzzi et al. (2011)
Pinguicula	Casper and Stimper (2009)
Polish vascular flora	Ivanova and Piekos-Mirkova (2003), Góralski et al. (2009) and Gacek et al. (2011)
Rubus	Thompson (1997)
Serapias	Bellusci and Aquaro (2008)
Slovak vascular flora	Marhold et al. (2007a, 2007b)
Veroniceae	Albach et al. (2008)

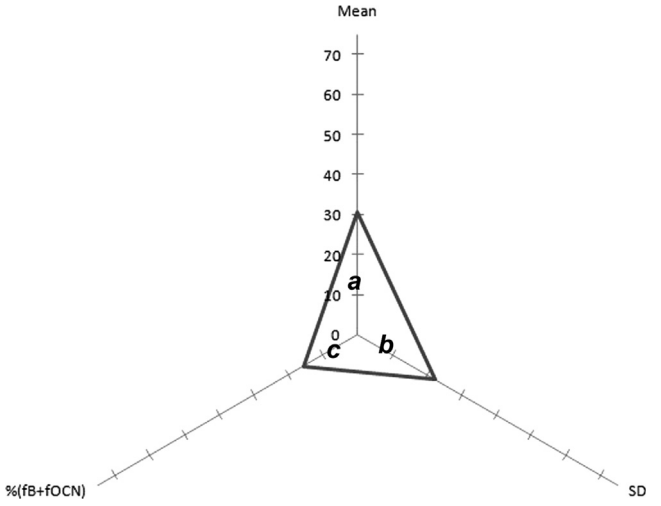


Fig. 1. Three-variables radar plot, where *a*=CN value, *b*=SD value, *c*=%(*f*B+*f*OCN) value.

3. Results

Table 2 shows the CN parameters calculated in our dataset. Mean CN ranged in animals from 14.68 (Cyclopoida) to 70.88